

Il catalogo connesso

Il catalogo ha subito negli anni una serie di profondi cambiamenti che lo hanno trasformato, pur mantenendo invece costanti alcuni elementi. Nel corso di questo intervento verranno ripercorse le tappe principali che hanno portato allo stato di definizione attuale di questo strumento, cercando di identificare le linee di tendenza di tali cambiamenti.

Il catalogo è uno strumento di mediazione, anzi, il principale strumento di mediazione tra le raccolte delle biblioteche e gli utenti per l'uso dei quali tali raccolte sono costruite. Esso costituisce una rappresentazione simbolica del patrimonio della biblioteca e rende conoscibile, recuperabile e utilizzabile il patrimonio stesso. Attraverso la sua struttura sindetica, tuttavia, il catalogo assolve anche a un secondo ordine di funzioni, quello di manifestare i collegamenti tra i documenti che la biblioteca seleziona, raccoglie e mette a disposizione. Questa duplicità di scopo e funzioni, riassumibile con i termini che spesso le definiscono nella letteratura professionale internazionale "to locate (finding function) vs. to colocate (gathering function)" è al centro di molte delle riflessioni che hanno fatto la storia della catalogazione bibliografica e che sono, tuttora, in parte non risolte.

A seguito dell'evoluzione tecnologica, in particolare a causa della creazione dell'infrastruttura di Internet e della nascita del Web, tale duplicità è cresciuta, assecondata da uno sdoppiamento tra le tipologie di utenti che utilizzano il catalogo: da una parte quella che definiamo utenza "remota", quella che non utilizzerà mai i servizi più direttamente collegati alla fisicità delle biblioteche, e dall'altra a quella "locale", quella cioè che invece, almeno potenzialmente, è in grado di utilizzare tali servizi. Rispetto a questa duplicità dell'utenza, Paul Weston individua anche due differenti modalità di utilizzo del catalogo, che pongono l'accento su alcune funzioni di tale strumento piuttosto che su altre. In particolare, l'utenza locale, con l'obiettivo dell'accesso diretto ai documenti di suo interesse, focalizzerà la sua attenzione sugli aspetti strettamente catalografici, quelli che riguardano la rappresentazione simbolica del patrimonio di una o più biblioteche, concentrandosi su dati non strettamente bibliografici, ma collegati all'uso locale delle risorse (es. dati di collocazione, di disponibilità per il prestito ecc.), ponendo l'accento sul catalogo come portale di accesso a specifici servizi. L'utenza remota, al contrario, sarà interessata quasi esclusivamente ai dati bibliografici, esaltando in un certo senso la funzione "repertoriale" del catalogo ma anche a un eventuale accesso diretto alla copia digitale, se esistente. In tal caso il catalogo tenderà a essere percepito come un portale di accesso all'informazione tout-court (Weston 2007).

Il tema dell'evoluzione del catalogo va affrontato, a mio avviso, secondo tre diversi filoni, o punti di vista, perché tre sono gli aspetti da considerare insieme:

1. La **struttura dell'informazione**, intesa come "scheme" dei metadati, o più in generale come la teoria e le tecniche della catalogazione;
2. Il risultato della catalogazione stessa, cioè i **dati**, cioè il valore che assumono i singoli attributi per ogni oggetto descritto nell'universo di riferimento;
3. L'infrastruttura e la **tecnologia** che permettono il funzionamento dell'interfaccia verso l'utente finale.

Nel corso degli ultimi anni si è verificata nella letteratura professionale un'accentuazione dell'importanza legata a questo ultimo elemento, probabilmente dovuta a una certa arretratezza tecnologica dello strumento OPAC rispetto a come contemporaneamente si stava modificando la tecnologia del web (il cd.

web 2.0), tanto che per un certo periodo, cinque o sei anni fa, si parlava di OPAC 2.0 (Marchitelli and Piazzini 2008). È invece importante avere uno sguardo a tutto tondo sui tre versanti, poiché altrimenti, i cambiamenti superficiali non investirebbero dal profondo funzioni e funzionamento del catalogo, ma rimarrebbero, come scrisse Roy Tennant, qualche anno fa, solo “rossetto sulle labbra del maiale” (Tennant 2007).

La teoria catalografica

Innanzitutto, bisogna interrogarsi sugli obiettivi, o funzioni, del catalogo, così come lo abbiamo definito finora.

L'avvio di quella che Michael Gorman indica, in un messaggio rivolto a Seymour Lubetzky in occasione del festeggiamento dei cento anni dell'illustre bibliotecario (Gorman 2000), come “Great Tradition” della catalogazione, e che arriva a comprendere Lubetzky (e in fondo Gorman stesso) si fissa al 1839 con la pubblicazione delle 91 regole di Antonio Panizzi per il catalogo a stampa del British Museum (edito poi a partire dal 1841) e prosegue con le regole redatte da Charles Coffin Jewett, pubblicate come draft nel 1852 e in veste definitiva nel 1853 (Jewett 1853), e che si sviluppa e si consolida con Charles Ammi Cutter, in particolare con la pubblicazione, nel 1875, delle sue Regole per un catalogo dizionario a stampa (Cutter 1875).

Le regole di Cutter sono particolarmente significative, poiché presentano la prima formalizzazione degli scopi del catalogo che, pur con qualche modifica, come vedremo oltre, riteniamo ancora validi. Secondo Cutter il catalogo deve “mettere una persona in grado di trovare un libro di cui sia noto: A) l'autore; B) o il titolo; C) o il soggetto; mostrare cosa la biblioteca possiede: D) di un determinato autore; E) su un determinato soggetto; F) di un determinato genere di scritti; facilitare la scelta di un libro: G) relativamente alla sua edizione (in senso bibliografico); H) relativamente al suo carattere (in senso letterario o topico)”.

Nel 1941 Andrew D. Osborn pubblica *The crisis in cataloging* (Osborn 1941), in cui evidenzia il caos e la contraddizione nei codici di catalogazione statunitensi editi fino ad allora. Da questo contributo parte la ricostruzione della teoria catalografica su basi moderne che annovera lo studio conosciuto poi come Report Henkle (Henkle 1946) e i contributi di Seymour Lubetzky (1953; 1960), di Eva Verona (1959) e, pur se in un filone a sé, gli studi di Shiyali Ramamrita Ranganathan.

Tali riflessioni, trascorse in poco più di un secolo, hanno il loro culmine nel 1961, con l'approvazione dei Principi di Parigi, pietra miliare nella storia della catalogazione, completati nel corso dell'incontro internazionale di esperti di catalogazione di Copenaghen del 1969, dalle cui risoluzioni scaturisce lo standard ISBD.

I principi di Parigi danno una nuova formalizzazione alle funzioni del catalogo già definite da Cutter. In essi infatti si prevede che “Il catalogo deve essere uno strumento efficace per accertare se la biblioteca contiene un libro particolare identificato per mezzo del suo autore e titolo, oppure se l'autore non è nominato nel libro, per mezzo del titolo soltanto, oppure se autore e titolo sono inadatti o insufficienti all'identificazione, un conveniente sostituto del titolo; quali opere di un particolare autore e quali edizioni di una particolare opera esistono nella biblioteca”.

Da questo momento e per quasi un trentennio, sulla riflessione filosofica intorno al catalogo prevarranno gli aspetti normativi. Sono gli anni nei quali nasceranno i maggiori codici di catalogazione nazionali, tra i quali il nostro (*Regole Italiane Di Catalogazione Per Autori* 1979).

Nel corso di questo periodo, si segnala la riflessione di Micheal Malinconico (Malinconico 1979), importante ai fini di questa ricostruzione, che propone una nuova declinazione delle funzioni del catalogo, che deve:

1. Consentire la localizzazione fisica di un particolare, ben determinato oggetto; 2. collegare le singole manifestazioni di un'opera particolare (le sue traduzioni e edizioni, comprese quelle in formati diversi); 3. collegare tutte le opere che sono il prodotto di un singolo agente responsabile della loro creazione, ovvero le opere che hanno una origine comune; collegare tutte le opere che trattano di un soggetto comune, ovvero le opere che hanno un'intenzione comune.

La formulazione proposta da Malinconico sottolinea ancora che le finalità del catalogo devono essere necessariamente almeno quelle due alle quali si accennava sopra: una attraverso la quale il catalogo individua certi documenti (locate, cfr. punto 1) e una nella quale il catalogo raggruppa, o collega, gli oggetti in base a caratteristiche prestabilite (colocate, cfr. punti 2-4).

L'interesse per la teoria catalografica torna poi più vivo con gli anni '90 del secolo scorso, quando l'IFLA promuove un gruppo di lavoro con lo scopo di verificare, ed eventualmente adeguare, i Principi di Parigi ai mutamenti di contesto intervenuti a partire da allora. Lo studio si conclude nel 1998 con la pubblicazione del rapporto finale su FRBR (IFLA Study Group on the Functional Requirements for Bibliographic Records et al. 1998). Il Rapporto si pone come modello concettuale, ad alto livello di astrattezza, che ha lo scopo di delineare "con l'uso di termini chiaramente definiti, le funzioni svolte da un record bibliografico rispetto ai vari media, alle varie applicazioni ed ai vari bisogni dell'utente" (IFLA Study Group on the Functional Requirements for Bibliographic Records et al. 1998, 7). Ai fini dell'analisi che stiamo qui conducendo, è proprio questa la parte più interessante dello studio: quella che porta a delineare le quattro funzioni alle quali il record bibliografico deve rispondere per soddisfare il bisogno informativo dell'utente:

- l'utente utilizza i dati per trovare materiali che corrispondono a determinati criteri, ad esempio per ricercare tutti i documenti su un determinato soggetto o ricercare una registrazione sonora edita con un titolo particolare;
- l'utente utilizza i dati reperiti per identificare un'entità, ad esempio per confermare che il documento descritto in una registrazione corrisponde al documento desiderato o per distinguere fra due testi o fra due registrazioni sonore che hanno lo stesso titolo;
- l'utente utilizza i dati per selezionare un'entità che corrisponda alle proprie necessità, ad esempio per selezionare un testo nella propria lingua o per scegliere una versione di software compatibile con l'hardware e con il sistema operativo a disposizione;
- l'utente utilizza i dati per acquisire od ottenere accesso all'entità descritta, ad esempio compilare l'ordine di acquisto di una pubblicazione, per inoltrare la richiesta di prestito di un libro, per accedere a un documento elettronico presente in un computer remoto.

È in base alle esigenze di trovare, identificare, selezionare e ottenere che FRBR definisce i requisiti funzionali (Guerrini 2000).

Più recentemente, a distanza di un decennio circa dalla pubblicazione di FRBR, vedono la luce i nuovi Principi internazionali di Catalogazione (ICP), risultato di otto anni di lavoro dell'IME ICC, IFLA Meeting of Experts on an International Cataloguing Code. L'idea di IME ICC nacque nel 2001, anno in cui ricorreva il quarantennale dei Principi di Parigi. Natalia N. Kasparova, dell'IFLA Cataloguing Section, propose la convocazione di un convegno internazionale per discutere i Principi alla luce del nuovo ambiente catalografico, profondamente mutato dal 1961 (Genetasio 2012).

Gli ICP, che pubblicati nel 2009, aggiornano ed estendono i Principi di Parigi, che riguardavano soltanto la scelta e la forma dell'intestazione, a ogni aspetto della catalogazione descrittiva; includendo indicazioni per le registrazioni di autorità e dichiarando l'intenzione di applicarsi anche alla catalogazione per soggetto. Estendono, per il discorso che ci riguarda più direttamente, le funzioni di localizzazione (trovare un libro di cui sia conosciuto l'autore, il titolo o il soggetto) e raggruppamento (trovare tutti i libri di un determinato autore o su un determinato soggetto o genere letterario) già presenti nei Principi di Parigi, aggiungendovi quelle definite da FRBR, in una prospettiva che enfatizza la centralità dell'utente. A tali obiettivi viene aggiunta, in linea con l'elaborazione di Elaine Svenonius (2000), la funzione di navigazione: "navigare in un catalogo e al di fuori [...] tramite la disposizione logica dei dati bibliografici e di autorità e la presentazione di chiare modalità per muoversi, compresa la presentazione delle relazioni fra le entità".

Un ulteriore elemento di cambiamento, forse ancora troppo recente per valutarne la portata storica ma che si annuncia tuttavia come fondamentale, è la comparsa nel 2010 di RDA che, nato come terzo aggiornamento del codice di catalogazione anglo americano (AACR), è passato rapidamente a rappresentare un elemento di novità rispetto ai suoi predecessori. Profondamente basato su FRBR e ICP, RDA si autodefinisce già a partire dal titolo come diverso. Scompare infatti il termine "catalogazione", sostituito dalla specificazione delle due componenti fondamentali della catalogazione stessa, la descrizione e la scelta dei punti di accesso, scompare il riferimento geografico, scompare anche l'idea del "codice". RDA è infatti «un insieme di linee guida e di istruzioni per la formulazione dei dati che consentano la scoperta delle risorse».

La crescente importanza del "dato" bibliografico

Veniamo quindi all'analisi dei cambiamenti che hanno coinvolto il dato bibliografico. La norma UNI-ISO 5127-5 del 1987 utilizza l'espressione "notizia bibliografica" ("Documentazione e Informazione. Vocabolario. Acquisizione, Identificazione e Analisi Di Documenti e Dati. UNI ISO 5127-5:1987") per tradurre "record" bibliografico.

L'oggetto così denominato è tipicamente espresso in un formato di marcatura specifico, come il MARC, nelle sue numerose varianti. Si tratta di un linguaggio concepito, alla fine degli anni '60, per agevolare lo scambio controllato di informazioni bibliografiche, rendendo disponibili ovunque, con rapidità e in forma accettata a livello internazionale, i dati bibliografici relativi a tutte le pubblicazioni di tutti i paesi.¹ Il MARC, basato sulla norma ISO 2709 ("Information and Documentation -- Format for Information Exchange. ISO 2709:2008") è evidentemente concepito per creare record da scambiare tra agenzie bibliografiche sul supporto di un nastro magnetico. Siamo nell'era del mainframe (fino all'inizio degli anni '80 non erano diffusi i microcomputer, tantomeno le reti, la cui diffusione si fa convenzionalmente risalire alla metà degli anni '90); l'informazione è registrata in forma lineare: il nastro su cui viene depositata ha un inizio, una fine, un verso di scorrimento. Il record presenta delle caratteristiche che permettono di utilizzarlo in quell'ambiente. È il caso, per esempio, della zona del record che chiamiamo "directory entry", o indice, che contiene in forma normalizzata l'elenco dei campi, con le dimensioni di ciascuno di essi indicate in numero di caratteri: è evidente che questo elemento risponda all'esigenza di separare e scorrere sul nastro i diversi record e, all'interno di ciascuno di essi, i singoli campi contenuti.

¹ Si tratta del programma IFLA UBC (Universal bibliographic Control) che nel 1987 venne, significativamente, fuso con altre due iniziative IFLA, UAP (Universal Availability Program) e IMP (International MARC Programme), nel progetto UBCIM (Universal Bibliographic control and International MARC Programme) (Su queste iniziative, si veda Solimine 1995).

Già nel 2002, però, anche se qualcuno l'aveva fatto anche prima, forse con meno rumore, si gridava alla morte del MARC (Tennant 2002a; Tennant 2002b). Se MARC, che pure si voleva morto, era tuttavia ancora in salute e veniva (e viene) ampiamente utilizzato per lo scambio dei dati bibliografici, esso non era utilizzato direttamente come formato di creazione e memorizzazione per i record bibliografici.

Nel back-office, la prima tappa dell'automazione è rappresentata dai database. In una base dati, tutti i metadati che prima erano contenuti nella scheda cartacea sono conservati in un file di computer . I metadati vengono separati in diversi attributi (o campi o celle) in modo regolare e tutti i record hanno lo stesso set (o sub - set) di attributi. Questa struttura è implicita nei metadati della scheda cartacea, redatta generalmente secondo ISBD, in cui gli attributi sono indicati da vari dispositivi di punteggiatura (come le parentesi) , ma non sempre identificati specificamente .

Questo tipo di file è noto come record flat , poiché tutti i metadati vengono memorizzati in una struttura bidimensionale , o piatta, monolitica (Dunsire 2009).

Il record piatto non è un modo efficiente di immagazzinare metadati, poiché obbliga la ripetizione di contenuti tra i diversi record , per esempio, come per esempio nomi di persone e organizzazioni. La ripetizione viene ridotta al minimo creando e memorizzando un singolo record, anche piatto, contenente i metadati per la persona o organizzazione, collegato poi al relativo record bibliografico tramite un identificatore numerico, una chiave. Lo stesso approccio funziona per i descrittori di soggetti tratti da vocabolari controllati: una sola registrazione di autorità per ogni soggetto è legata a molte descrizioni bibliografiche .

Questo tipo di struttura informativa è nota come database relazionale.

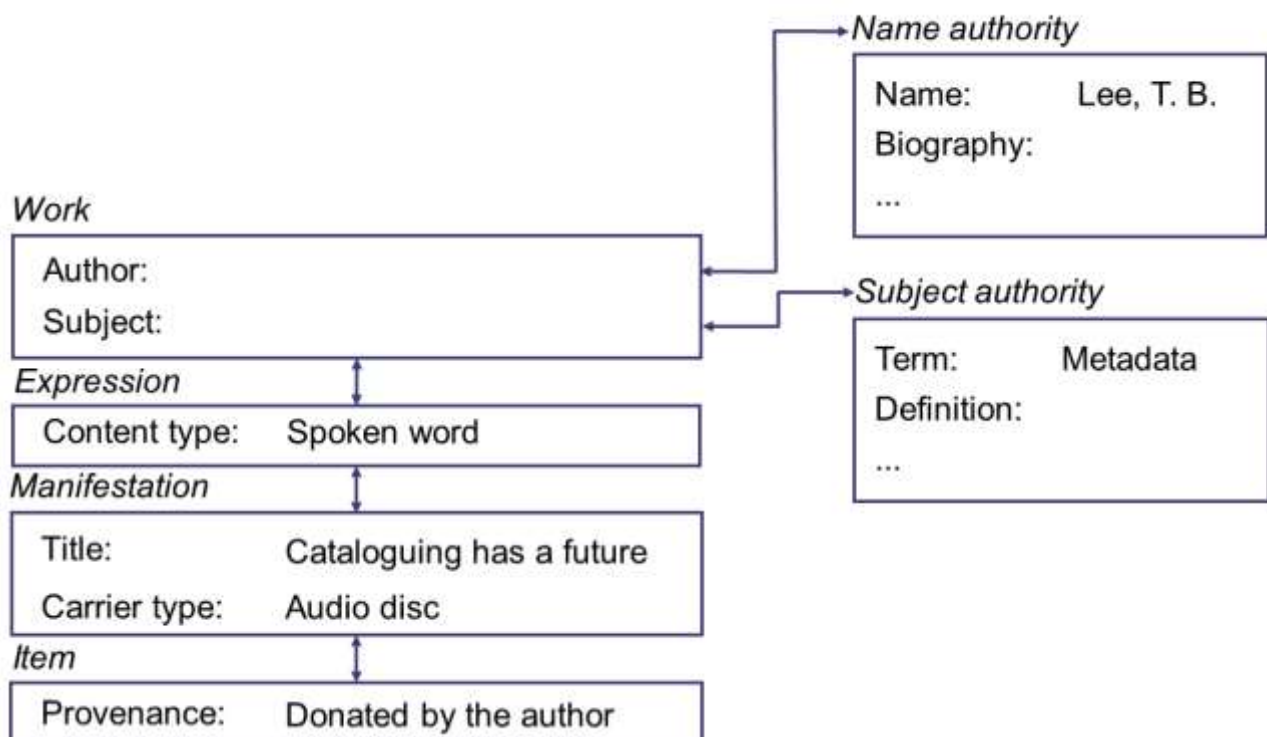


Figura 1: Rappresentazione di record in una database relazionale

Questo metodo di memorizzazione di metadati descrittivi, separati dalle voci di autorità che fungono da punti di accesso per la descrizione, ha cominciato a essere attuato nel 1970. Esso è tuttora utilizzato dalla maggior parte dei cataloghi automatizzati.

Ancora FRBR, nel 1998, fa riferimento, e proprio nel suo titolo, al record così concepito, anche se nella sostanza il modello FRBR innova profondamente la costruzione della struttura dei dati, fino ad allora comunque bidimensionale. FRBR raggruppa gli attributi delle entità in quattro diversi insiemi, collegati in un modo più o meno gerarchico. Al livello più basso è l'esemplare, contenente i metadati di una singola copia di una risorsa. In questo esempio, la nota relativa alla provenienza applica esclusivamente alla specifica copia specifica.

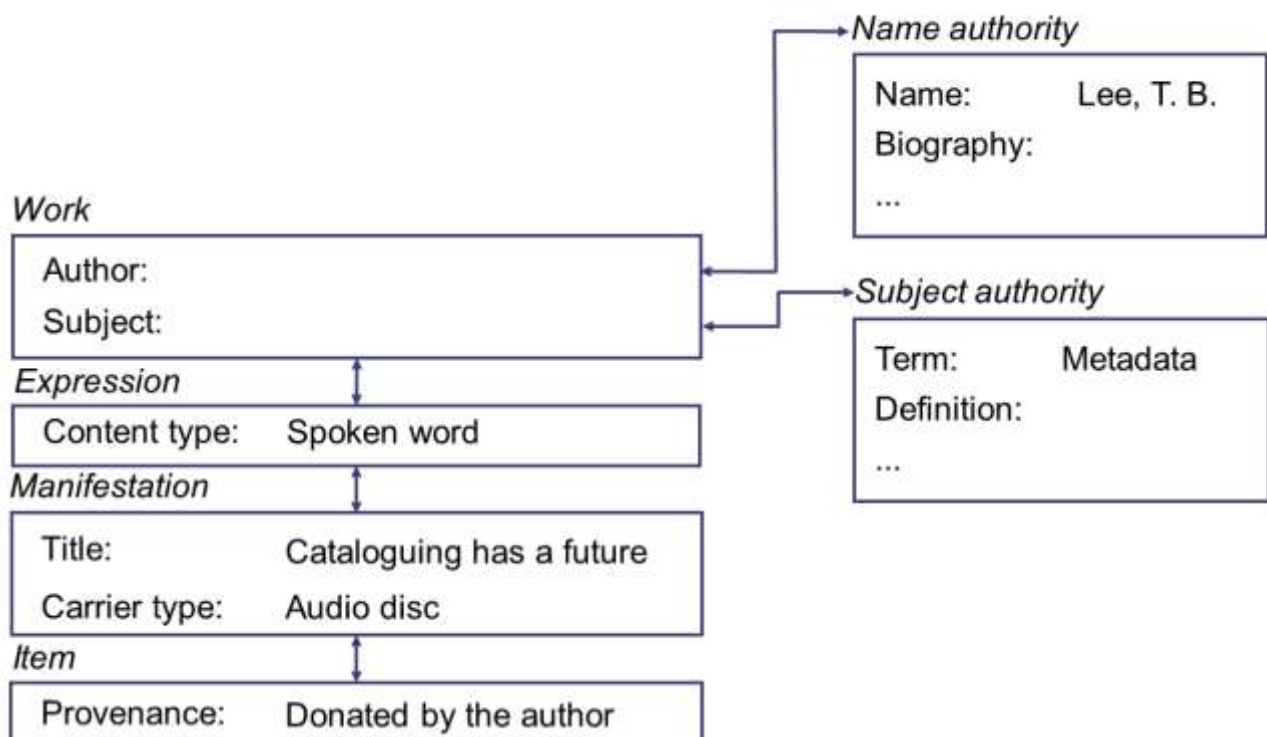


Figura 2: Rappresentazione di un record multidimensionale FRBR

Il successivo raggruppamento di metadati è la Manifestazione, che descrive il modo in cui l'informazione è veicolata dalla risorsa. I metadati includono qui, per esempio, il titolo preso dall'etichetta del disco, oltre che l'indicazione del tipo di supporto.

Continuando a salire nella gerarchia, gli altri due gruppi sono Espressione e Opera. I metadati dell'Espressione descrivono il contenuto della risorsa, mentre quelli dell'Opera descrivono il concetto astratto che le sta dietro la risorsa, per esempio l'autore o il soggetto.

Un'opera può avere più di un'espressione (ad esempio una traduzione), che può avere più di una manifestazione (ad esempio, audiocassette e compact disc), che può avere più di un esemplare (copia). Ciò si traduce in una duplicazione, poiché due espressioni utilizzano i metadati della medesima Opera. Applicando i medesimi principi per la costruzione di database descritti sopra per le intestazioni di autorità, un record FRBR deve quindi essere conservato in quattro file separati collegati tramite una sorta di numero di identificazione o di controllo.

Negli anni successivi alla pubblicazione di FRBR si indebolisce ulteriormente l'idea di "record", che viene pian piano sostituita dal termine "dato", anche nei titoli di importanti documenti, come in IFLA ICP (*Statement of International Cataloguing Principles (ICP) and Its Glossary* 2009) o in FRAD (Patton et al. 2009).

Questo permette di porre l'accento sulla frammentazione dell'informazione, ricomposta attraverso il linking (dal record monoblocco al mosaico dei dati), tanto è vero che qualcuno ha parlato di un cambiamento, dal "cataloguing" al "catalinking".²

In questo esempio, i metadati del componente Opera non contengono alcuno specifico valore, solo la struttura in forma di nomi di attributi e collegamenti al contenuto dei file autorità.

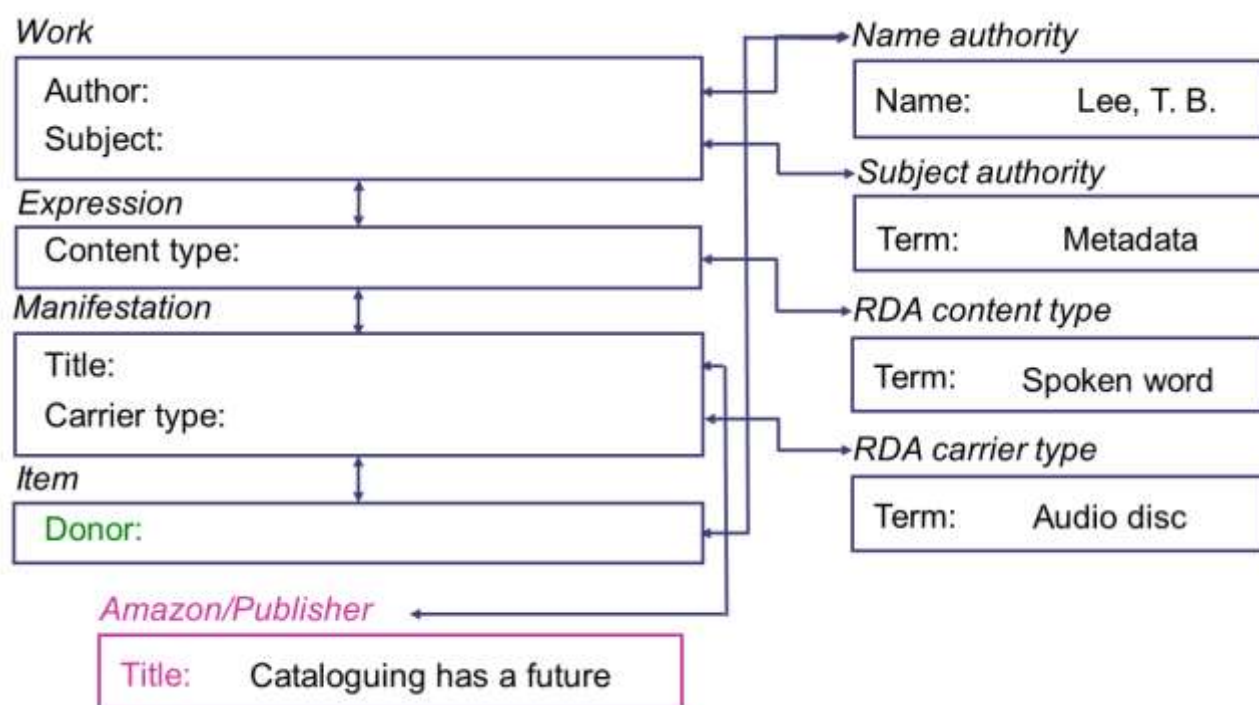


Figura 3: Dati, frammentati e collegati

L'attributo provenienza nel componente "item" si riferisce all'autore inteso come donatore ed è una duplicazione implicita di una parte del contenuto dell'attributo dell'autore, sottoposta al controllo di autorità. Per minimizzare tale duplicazione si raffina l'attributo "provenienza" nell'attributo "donatore", più specifico, e si isola il riferimento all'autore come il contenuto di tale attributo. L'attributo donatore viene quindi collegato alla medesimo dato di autorità utilizzato come contenuto dell'attributo "autore".

Tutto questo collegare dati lascia il solo titolo della manifestazione all'interno di quello che originariamente era il record bibliografico. Ma si potrebbe collegare l'attributo con l'archivio di un editore o di una libreria... per arrivare a un record-scheletro, che contenga solo i nomi degli attributi e le chiavi per i collegamenti. Il record è così ridotto alla sua struttura, ed effettivamente è estinto.

Un'altra spinta assai forte al cambiamento, nel senso della disgregazione del record, è venuta dall'avvento del web semantico e dei Linked Open Data, tema approfondito da altri interventi di questa sessione, come

² Roy Tennant, ad ALIA online 2013, <http://storify.com/ALIAonline/keynote-roy-tennant-jon-voss-and-ingrid-mason>

quello di Antonella Iacono, sul quale quindi ora sorvolo, per arrivare a una particolare applicazione del web semantico ai dati bibliografici, quale si sta delineando BibFrame.

Un punto di svolta nella storia recente dei formati bibliografici è infatti rappresentato dal rapporto “On the record”, pubblicato da un gruppo di studi costituito dalla Library of Congress per ragionare sul futuro del record bibliografico e sull’uso del MARC (Library of Congress Working Group on the Future of Bibliographic Control 2008). Il rapporto conteneva il suggerimento per la Library of Congress di iniziare a lavorare, anche coinvolgendo altre comunità oltre a quella bibliotecaria, per specificare e implementare un formato di scambio di dati bibliografici. Era il momento di smettere di discutere di possibili alternative al MARC e iniziare concretamente a lavorare a tale cambiamento.

Nel novembre 2012, viene rilasciata la prima versione del modello (US Library of Congress and Zepheira). Senza sorprese, alla luce della letteratura dell’ultimo decennio (Kroeger 2013), BibFrame è flessibile ed estensibile, basato su RDF e codificato in XML. Più sorpresa desta il fatto che il modello semplifichi l’universo rappresentato in RDF, riducendo le quattro entità del primo gruppo FRBR (Opera, Espressione, Manifestazione e Item) a due solamente, l’Opera e l’Istanza (che grossomodo possiamo far coincidere con Espressione + Manifestazione di FRBR), con l’aggiunta dell’Annotazione, che equivale all’esemplare ma può comprendere anche “Other linked data stuff”, come l’immagine della coperta, o le recensioni.

BibFrame è ancora costituito da una serie di draft, tool sperimentali e la discussione è ancora aperta. Possiamo tuttavia dire, a questo punto con qualche certezza, che in futuro i record bibliografici saranno disgregati al momento della loro creazione, nasceranno in momenti e da agenti probabilmente diversi tra di loro, ma verranno riaggregati e presentati unitariamente agli utenti, anzi, probabilmente, saranno gli utenti a definire la presentazione dei dati bibliografici secondo i loro scopi.

Questa è la realizzazione concreta del principio della navigabilità (nel catalogo e oltre) postulata da ICP.

Tecnologia

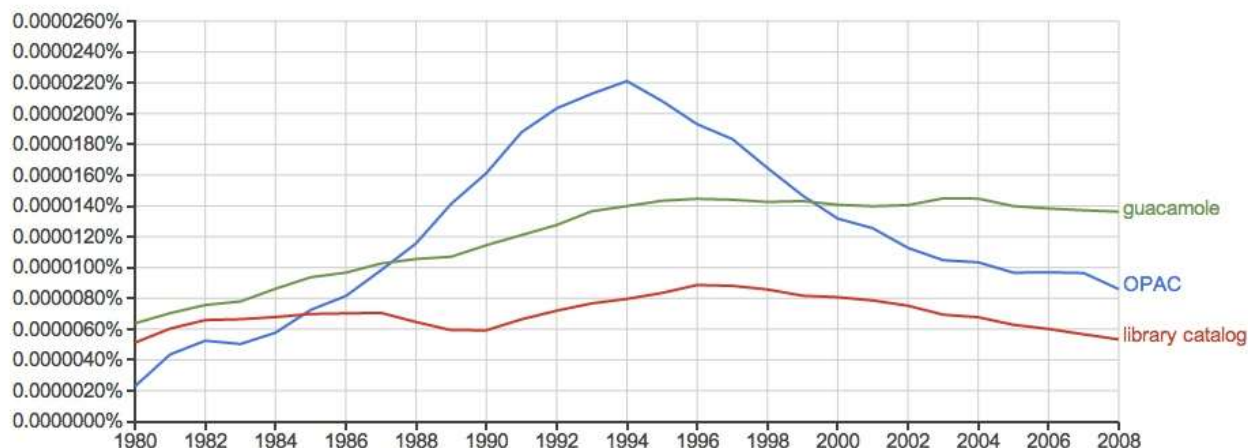
Sin qui abbiamo visto come, nel corso dell’ultimo secolo e mezzo, la teoria catalografica abbia ampliato le funzioni del catalogo a comprendere in maniera esplicita anche la navigazione all’interno e all’esterno del catalogo. Allo stesso modo, abbiamo seguito la trasformazione del dato bibliografico, che nell’ultimo cinquantennio è passato dalla forma del record monolitico, contenente tutte le informazioni necessarie alla comprensione del contenuto del record stesso, a un man mano sempre più disaggregato insieme di informazioni descrittive e relazioni tra entità, che a loro volta si stanno trasformando sempre più in collegamenti su web, con valore sempre più semantico.

Veniamo quindi all’aspetto tecnologico, quello degli applicativi software che veicolano la presentazione di questi dati all’utenza finale delle biblioteche, qualsiasi essa sia.

Dal 1984, a partire da un ormai celebre articolo (Hildreth 1984), si usa individuare diverse tappe nella storia dello strumento di ricerca nei cataloghi di biblioteca che siamo ormai abituati a definire “OPAC”, che nacque all’inizio degli anni ’70. Si fa risalire l’origine del nome, acronimo di Online Public Access Catalogue, al primo catalogo online delle università dell’Ohio, probabilmente il primo catalogo online al mondo, realizzato nel 1971.

Si noti incidentalmente che il termine OPAC, di uso ormai comune tra gli addetti ai lavori, ma decisamente meno tra gli utenti finali, fa riferimento al sistema di ricerca tipico degli applicativi di gestione integrata di

biblioteca (ILS) e quindi è, nel contesto attuale, applicato in maniera forzata alle interfacce di ricerca, che sono generalmente separate e indipendenti dal sistema di gestione del backoffice. Questo aspetto è evidente anche nelle ricerche condotte su web per tale termine, in declino nell'ultimo decennio, secondo quanto afferma Roy Tennant (2014), che ne propone quindi con forza l'abbandono.



A partire dall'inizio dell'epoca dell'automazione del catalogo, e in particolare della disponibilità pubblica di interfacce di ricerca, si afferma, dall'inizio degli anni '70, quella che Hildreth definisce "prima generazione". Si tratta di strumenti di information retrieval semplici, che permettono solo poche tipologie di ricerca differenti, sostanzialmente reiterando il modello del catalogo cartaceo, con ridotti canali di accesso, quelli definiti da ICP del 1961, autore e titoli soltanto. Anche la presentazione dei dati ripete le forme della scheda cartacea e non sfrutta le potenzialità del mezzo digitale attraverso il quale i record sono diffusi.

La seconda generazione di OPAC risente del maggiore sviluppo dei database di ricerca commerciali e inizia a diffondersi a partire dagli anni '80. Lo scopo che si prefiggono gli sviluppatori di questi strumenti, e che in gran parte raggiunsero, è quello di coniugare la familiarità degli utenti con la fisionomia del catalogo cartaceo, sfruttando però più a fondo le tecnologie dell'information retrieval elettronico, per esempio ampliando il numero dei canali di ricerca, permettendo la ricerca per parole contenute e non solo per parte iniziale o frase, o creando la possibilità di raffinare le ricerche senza ripetere ex-novo la query.

A questa, con gli anni '90 succede una terza generazione, quella che Hildreth chiama degli OPAC estesi, o meglio E³OPAC, dove questa "E al cubo" rappresenta le tre caratteristiche principali ravvisate dallo studioso in tali strumenti:

- Enhanced: per funzionalità e usabilità
- Expanded: Indici, record, collezioni (diventando cioè uno strumento di accesso all'intera collezione della biblioteca)
- Extended: attraverso link, reti, e gateway verso altre collezioni, sistemi informativi e risorse.

Tali strumenti informatici, che come quelli appartenenti alla fine della generazione precedente, si presentano con un'interfaccia utente grafica (Graphical User Interface, GUI) e non più a caratteri si sono mantenuti stabili sul mercato a lungo, almeno fino a quando, dopo la metà degli anni Duemila inizia a manifestarsi una forte insoddisfazione per gli OPAC, rappresentata in decine di post su blog (si veda per tutti la "trilogia" di Schneider 2006a; 2006b; 2006c), fino ad arrivare addirittura a una canzone con tanto di videoclip (Smith 2006) che canta che "gli OPAC fanno schifo".

Sono gli anni in cui il web ha subito alcuni cambiamenti, dal punto di vista tecnologico e funzionale, che vanno sotto il nome di "Web 2.0" (per una disamina complessiva del fenomeno, v. Marchitelli 2009). Tali innovazioni, unite alla tendenza alla "dis-integrazione degli applicativi gestionali" (K. Antelman, cit. in Weston 2008) si sono poi riverberate anche negli strumenti di ricerca sul catalogo, ormai separati dal resto dell'ILS. Tali strumenti, che potremmo definire di quarta generazione, sono stati rapidamente etichettati proprio "cataloghi di nuova generazione" (NGC, New Generation Catalogues), o discovery tool (Marchitelli 2013; Marchitelli and Frigimelica 2012).

Si è tentato quindi di porre rimedio alle difficoltà incontrate dagli utenti rendendo gli strumenti di ricerca per i cataloghi di biblioteca sempre più "Google-like". Questo ha significato una serie di cambiamenti strutturali all'architettura dell'informazione fornita dai cataloghi e una serie nutrita di nuove funzionalità che facilitano il compito dell'utente nel momento in cui si mette alla ricerca di informazioni.

In particolare, tali cambiamenti sono stati indirizzati:

1. alla ottimizzazione e semplificazione delle modalità di funzionamento rese disponibili agli utenti, con l'introduzione della box di ricerca singola in home page;
2. alla "sfocatura" della precisione delle ricerche attraverso l'uso della logica fuzzy e altri stratagemmi come i suggerimenti in merito ai termini digitati per la ricerca (es. "Forse cercavi:") e alle possibilità di allargare il campo di ricerca, nell'ottica di fornire sempre all'utente un qualche risultato per qualsiasi ricerca effettuata, come nel comportamento tipico di un motore di ricerca web;
3. alla gestione di un ordinamento dei risultati basato sulla "rilevanza", in aggiunta ai tradizionali alfabetico e cronologico;
4. alla semplificazione dell'analisi dei risultati forniti attraverso il raggruppamento dinamico dei risultati di ricerca (navigazione "a faccette");
5. a una visualizzazione dei dati più ricca, compresi anche a quelli non prettamente bibliografici e i contenuti generati dagli utenti (UGC, User Generated Contents) esito in parte dell'arricchimento dei cataloghi già in atto da almeno un ventennio, e in parte portato dal fenomeno che ha preso il nome di Web "2.0".

Tali discovery tool riscosero da subito successo, in particolare grazie alle interfacce rinnovate e a una usabilità analoga a quella degli applicativi che su web avevano grande fortuna (come Google, Amazon o Facebook, per citare solo i più rappresentativi in tal senso).

Uno dei cambiamenti forse più importanti, probabilmente quello più gravido di conseguenze, però, era decisamente meno visibile. Faccio riferimento alla trasformazione degli algoritmi di ricerca, che anche per gli OPAC introdusse allora una logica di tipo probabilistico ("fuzzy") in aggiunta a quella booleana sulla quale avevano basato fino a quel momento le loro funzionalità di information retrieval.

Sono state compiute nel corso di questi ultimi anni diverse ricerche che hanno cercato di individuare le caratteristiche fondamentali di un catalogo di nuova generazione, proponendo poi un benchmarking tra i diversi software. Una delle più interessanti, anche se non troppo recente, è stata pubblicata nel 2012 su Library Hi Tech.

Particolarmente utile la "checklist" individuata dagli autori, che comprende:

1. Un unico punto di accesso a tutte le informazioni biblioteca. Il catalogo della biblioteca dovrebbe mettere a disposizione una sola interfaccia per la ricerca di tutti i materiali della biblioteca, compresi i puntatori agli articoli in banche dati elettroniche, nonché le registrazioni di libri e di collezioni digitali. Una ricerca dovrebbe recuperare insieme tutti i materiali pertinenti
2. Un'interfaccia che corrisponda allo stato dell'arte. I cataloghi di biblioteca dovrebbe avere un design moderno simile a quello di siti web commerciali, di e-business. Questo criterio è molto soggettivo e come tale è difficile da quantificare. I cataloghi di nuova generazione dovrebbero comunque ispirarsi a siti noti e utilizzati come Google o Amazon
3. Contenuti arricchiti; i cataloghi dovrebbero includere le immagini delle copertine dei libri, contenuti generati dagli utenti, come per esempio commenti, descrizioni, classificazioni e tag cloud, a differenza dei cataloghi tradizionali, dove tutti i punti di accesso alle informazioni sono creati da bibliotecari e catalogatori professionisti
4. La navigazione a faccette. Questi applicativi dovrebbero essere in grado di visualizzare i risultati della ricerca come insiemi di categorie sulla base di alcuni criteri, come date, linguaggi, formati, disponibilità, luoghi, utenti etc. L'obiettivo è quello di condurre una ricerca iniziale molto semplice, ricerca utilizzando parole chiave e poi perfezionare i risultati selezionando faccette che si saranno determinate
5. Casella di ricerca semplice, per parola chiave, presente in ogni pagina. Il catalogo di nuova generazione si presenta con una finestra di ricerca semplice per parole chiave che assomiglia a quella di Google o Amazon. Dovrebbe tuttavia essere disponibile anche un link alla ricerca avanzata. La casella di ricerca semplice, dovrebbe apparire in ogni pagina dell'interfaccia utente. Sebbene questa funzionalità sia considerata una delle caratteristiche più importanti in questa nuova generazione di strumenti, in realtà essa non è implementata in maniera diffusa. L'indagine citata mostra che la maggior parte dei cataloghi di biblioteca ancora oggi non offre una ricerca "google-like" come pagina iniziale, poiché generalmente i bibliotecari preferiscono la ricerca avanzata, ritenendo che la ricerca semplice produca risultati con minor precisione
6. Rilevanza. Una delle maggiori critiche che arriva da parte dei bibliotecari riguarda proprio l'ordinamento per rilevanza, che spesso è determinata secondo algoritmi non trasparenti. Inoltre, le statistiche di circolazione dovrebbero influenzare l'ordinamento dei risultati: i documenti più acceduti dovrebbero essere visualizzati più in alto nella lista dei risultati; altro elemento da considerare può essere la quantità di copie a disposizione, maggior numero di copie, maggiore importanza del risultato corrispondente
7. Forse cercavi ...? Un meccanismo di controllo della digitazione dovrebbe essere presente nei cataloghi di nuova generazione. Quando si verifica un errore nella ricerca, dovrebbe comparire una segnalazione dell'ortografia corretta, cliccando sulla quale si esegue una ricerca. Anche questo meccanismo, tuttavia, deve essere chiaramente identificato e non sostituirsi all'azione autonoma e libera dell'utente
8. Suggerimenti e materiali correlati. Pratica comune nei siti di e-commerce, al cliente vengono mostrati ulteriori elementi con un suggerimento del tipo "I clienti che hanno acquistato questo prodotto hanno comprato anche...". Allo stesso modo i cataloghi di nuova generazione dovrebbero suggerire libri ai lettori sulla base dei log delle transazioni, nella forma di "I lettori che hanno preso in prestito questo libro hanno preso in prestito il seguente..." oppure di un link a "Letture consigliate"

9. Contributi degli utenti. Questi strumenti consentono agli utenti di aggiungere dati ai vari record bibliografici. Gli utenti possono inserire descrizioni, riassunti, recensioni, critiche, commenti, voti e ranking, tagging o folksonomie. Gli utenti di oggi fanno sempre maggiore riferimento a ciò che altri utenti hanno da dire su articoli trovati online e valutano con attenzione i pareri di quelli che sentono essere loro pari. Una nuvola di tag può fornire inoltre punti di accesso e parole chiave descrittive utilizzate di frequente

10. Feed RSS, che permettono agli utenti di connettersi ai contenuti di loro interesse per aggiornamenti frequenti. I cataloghi di nuova generazione includono feed RSS in modo che gli utenti possano ricevere elenchi di nuove accessioni, dei libri più prestati o consultati, ricerche salvate e anche connettere al catalogo il proprio blog o pagina del feed reader

11. L'integrazione con i siti di social network. Quando catalogo di una biblioteca è integrato con i siti di social network, gli utenti possono condividere link alle risorse della biblioteca con i loro amici su social network come Twitter e Facebook

12. Link persistenti. URL stabili possono essere copiate e incollate e servire da collegamento permanente a questo record; facilita inoltre la scansione da parte dei *crawler* dei motori di ricerca. Anche questa è una caratteristica importante dei cataloghi di nuova generazione.

Questa assimilazione dei cataloghi ai motori di ricerca, tuttavia, ha implicitamente imposto anche un cambiamento del paradigma di funzionamento di tali strumenti, che ormai si riflette anche nel nome, discovery tool: si è infatti passati infatti dal modello “ricerca”, proprio degli strumenti di information retrieval” al modello “scoperta”.

La logica booleana, infatti, ha connotato in sé l'alternativa dualistica di Vero e Falso (0 e 1, nell'applicazione binaria che così tanta fortuna avrà in ambito informatico). Tale logica è basata su certezze, tanto è vero che le proposizioni booleane permettono di costruire delle “tavole di verità”, mentre la *fuzzy logic* consente di determinare una scala di eventi, che hanno una maggiore o minore possibilità di verificarsi, andando a costituire una scala di valori decimali, intermedi tra lo zero e l'uno.

Questo passaggio, che in qualche modo è paradigmatico, trasforma la certezza in probabilità, l'uguaglianza in somiglianza, la precisione della risposta in un ranking delle risposte, il retrieval di informazione già nota in discovery di informazione ignota.

Condurre una ricerca, significa avere già in mente un documento o un contenuto informativo che si vuole recuperare. La ricerca è condotta formulando una query in maniera da aumentare le possibilità di creare una corrispondenza con il contenuto che si desidera. In una query concepita a questo scopo, i termini utilizzati tendono a essere piuttosto descrittivi, in modo da qualificare esattamente il bisogno informativo, lasciando poco o nessun margine all'ambiguità. Il percorso di scoperta, invece, ha una natura più esplorativa, spesso guidata da un bisogno generico, o comunque meno esplicitamente dichiarato. La query inserita si trasforma quindi in un punto di partenza per un'esplorazione, mirata ad apprendere qualcosa su un determinato argomento. Come, nel web, i link ipertestuali facilitano un rapido spostamento tra temi e informazioni correlati, così un discovery tool correda i risultati di una serie di faccette, utilizzabili come link. Tali collegamenti rappresentano le differenti dimensioni dei risultati ottenuti e permettono di spostarsi all'interno dell'ambito informativo che si è arrivati a delimitare.

Gli strumenti per la scoperta di contenuti, cioè i discovery tool, offrono dunque un approccio diverso all'informazione. Essi mirano a fornire consapevolezza, piuttosto che una risposta specifica, aggregando

contenuti relativi a un particolare ambito di interesse. L'utente si muove attraverso l'informazione correlata a un certo argomento per avere la consapevolezza di ciò che sta accadendo; in sostanza un discovery tool nasce con l'obiettivo di fornire la certezza di non mancare le informazioni più importanti rispetto a una certa tematica. Offrono dunque informazioni inattese, perché gli utenti non avranno necessariamente lo specifico risultato che stanno cercando, quanto piuttosto, in maniera serendipitaria, notizie riguardo al tema di loro interesse in un delimitato periodo di tempo.

Il percorso fin qui seguito, partito dalla metà dell'Ottocento per quanto riguarda la teoria catalografica, e dalla metà dello scorso secolo per quanto riguarda la registrazione dei dati su supporto informatico e le tecnologie che la rendono possibile, è arrivato a tratteggiare l'evoluzione del catalogo su tre distinti filoni. Siamo però probabilmente alle soglie di altri importanti cambiamenti, a partire dalla base che abbiamo ricostruito fin qui. Qualsiasi analisi, dunque, è ora destinata a invecchiare presto.

Riferimenti bibliografici

- Cutter, Charles A. 1875. *Rules for a Printed Dictionary Catalogue*. Washington, DC: The Library of Congress. <https://archive.org/details/cu31924029518978>.
- "Documentazione e Informazione. Vocabolario. Acquisizione, Identificazione e Analisi Di Documenti e Dati. UNI ISO 5127-5:1987."
- Dunsire, Gordon. 2009. "A Short History of the Evolution of the Catalogue Record." <http://www.gordondunsire.com/pubs/pres/EvolCatRec.ppt>.
- Genetasio, Giuliano. 2012. "The International Cataloguing Principles and their future." *JLIS.it* 3 (1). doi:10.4403/jlis.it-4936. <http://leo.cilea.it/index.php/jlis/article/view/4936>.
- Gorman, Michael. 2000. "Seymour Lubetzky Uomo Di Principi." Translated by Agnese Galeffi. *Biblioteche Oggi* 28 (6): 6–11.
- Guerrini, Mauro. 2000. "Le Funzioni Del Catalogo dall'ICCP a FRBR". AIB-WEB. <http://www.aib.it/aib/sezioni/toscana/conf/frbr/guerrini.htm>.
- Henkle, Herman H., ed. 1946. *Studies of Descriptive Cataloging. A Report to the Librarian of Congress by the Director of the Processing Department ...* Washington: U.S. Govt. Print Off. <http://catalog.hathitrust.org/Record/001162946>.
- Hildreth, Charles R. 1984. "Pursuing the Ideal: Generations of Online Catalogs." In *Online Catalogs, Online Reference: Converging Trends : Proceedings of a Library and Information Technology Association Preconference Institute, June 23-24, 1983, Los Angeles*, edited by Brian Aveney, Brett Butler, and American Library Association Library and Information Technology Association (U.S.). Chicago: American Library Association.
- IFLA Study Group on the Functional Requirements for Bibliographic Records, International Federation of Library Associations and Institutions, Section on Cataloguing, and Standing Committee. 1998. *Functional Requirements for Bibliographic Records Final Report*. München: K.G. Saur. <http://catalog.hathitrust.org/api/volumes/oclc/39964684.html>.
- "Information and Documentation -- Format for Information Exchange. ISO 2709:2008."
- Jewett, Charles C. 1853. *On the Construction of Catalogues of Libraries, and Their Publication by Means of Separate, Stereotyped Titles, with Rules and Examples*. Washington, DC: The Smithsonian Institution. <https://archive.org/details/constcatlib00jewerich>.
- Kroeger, Angela. 2013. "The Road to BIBFRAME: The Evolution of the Idea of Bibliographic Transition into a Post-MARC Future." *Cataloging & Classification Quarterly* 51 (8): 873–90. doi:10.1080/01639374.2013.823584.
- Library of Congress Working Group on the Future of Bibliographic Control. 2008. *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control*. Washington, DC: Library of Congress. <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>.

- Lubetzky, Seymour. 1953. *Cataloging Rules and Principles: a Critique of the A.L.A. Rules for Entry and a Proposed Design for Their Revision*. Washington, DC: Processing Dept., Library of Congress.
- . 1960. *Code of Cataloging Rules; Author and Title Entry*. Chicago, IL: American Library Association.
- Malinconico, S. Michael. 1979. *The Future of the Catalog: The Library's Choices*. Professional Librarian Series. White Plains, N.Y: Knowledge Industry Publications.
- Marchitelli, Andrea. 2009. "Stat Rosa Pristina Nomine? Biblioteche e Cataloghi Ai Tempi Del Nuovo Web." *AIDAlnformazioni : Rivista Di Scienze Dell'informazione* 27 (3-4): 31 – 43.
- . 2013. "Gli OPAC Di Nuova Generazione e i Discovery Tool." In *Rapporto Sulle Biblioteche Italiane 2011-2012*. Roma: AIB.
- Marchitelli, Andrea, and Giovanna Frigimelica. 2012. *OPAC*. ET 29. Roma: Associazione Italiana Biblioteche. <http://www.aib.it/negozio-aib/et/opac/>.
- Marchitelli, Andrea, and Tessa Piazzini. 2008. "OPAC, SOPAC e Social Networking: Cataloghi Di Biblioteca 2.0?" *Biblioteche Oggi* 26 (2): 82–92.
- Osborn, Andrew D. 1941. "The Crisis in Cataloging." *The Library Quarterly* 11 (4): 393–411.
- Patton, Glenn E, International Federation of Library Associations and Institutions IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), Cataloguing Section, Standing Committee, International Federation of Library Associations and Institutions, Section on Classification and Indexing, and Standing Committee. 2009. *Functional Requirements for Authority Data a Conceptual Model*. München: K.G. Saur. <http://public.eblib.com/EBLPublic/PublicView.do?ptilID=453934>.
- Regole Italiane Di Catalogazione Per Autori*. 1979. Roma: Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche.
- Schneider, Karen G. 2006a. "How OPACs Suck, Part 3: The Big Picture." *ALA Tech Source*. <http://www.alatechsource.org/blog/2006/05/how-opacs-suck-part-3-the-big-picture.html>.
- . 2006b. "How OPACs Suck, Part 1: Relevance Rank (Or the Lack of It)." *ALA Tech Source*. <http://www.alatechsource.org/blog/2006/03/how-opacs-suck-part-1-relevance-rank-or-the-lack-of-it.html>.
- . 2006c. "How OPACs Suck, Part 2: The Checklist of Shame." *ALA Tech Source*. <http://www.alatechsource.org/blog/2006/04/how-opacs-suck-part-2-the-checklist-of-shame.html>.
- Smith, Brian. 2006. "New Song: 'The OPAC Sucks'." *The Laughing Librarian*. http://www.laughinglibrarian.com/2006_11_01_archive.html.
- Solimine, Giovanni. 1995. *Controllo Bibliografico Universale*. ET 8. Roma: Associazione italiana biblioteche : Editrice Bibliografica.
- Statement of International Cataloguing Principles (ICP) and Its Glossary*. 2009. Saur.
- Svenonius, Elaine. 2000. *The Intellectual Foundation of Information Organization*. MIT.
- Tennant, Roy. 2002a. "MARC Must Die!" *Library Journal* 127 (17): 26–28.
- . 2002b. "MARC Exit Strategies" 127 (19): 27.
- . 2007. "Lipstick on a Pig 2.0." *The Digital Shift*. <http://www.thedigitalshift.com/2007/05/roy-tennant-digital-libraries/lipstick-on-a-pig-2-0/>.
- . 2014. "The OPAC Is Dead." *The Digital Shift*. Accessed March 4. <http://www.thedigitalshift.com/2014/02/roy-tennant-digital-libraries/the-opac-is-dead/>.
- US Library of Congress, and Zepheira. "BIBFRAME Model Overview." <http://bibframe.org>.
- Verona, Eva. 1959. "Literary Unit Versus Bibliographical Unit." *Libri* 9 (2): 79–104.
- Weston, Paul Gabriele. 2007. "La gestione elettronica delle biblioteche." In *Biblioteconomia: principi e questioni*, edited by Giovanni Solimine and Paul Gabriele Weston, 221–56. Roma: Carocci.
- . 2008. "Caratteristiche degli opac e strategie delle biblioteche." *Biblio* 11 (1). <http://www.aib.it/aib/sezioni/emr/bibtime/num-xi-1/weston.htm>.